

ABSTRACT OF THE DISCLOSURE

[0048] A system and method of identifying unique email messages in a large scale enterprise environment using an external server and a database system. Message uniqueness is determined by assigning a message tag to each message based on properties of the email message. The message tag may be computed using a hashing algorithm to speed indexing and comparisons. The message tag is compared with an index file of message tags associated with pre-existing email messages. If a matching message tag is found in the index file, the email message is not unique. Otherwise, the email message is unique and the message tag is added to the index file. The system may include a relational database for storing the index file. An archiving system and method using the uniqueness checking feature of the present invention are also disclosed.